(2)

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

AN EMPIRICAL ANALYSIS OF ENLISTMENT INTENTIONS
AND SUBSEQUENT ENLISTMENT BEHAVIOR

by

Richard Paul Snyder

September, 1990

Thesis Co-Advisors: Linda Gorman
George W. Thomas

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School — 6b. OFFICE SYMBOL OR	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c. ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000	7b. ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION — 8b. OFFICE SYMBOL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. — PROJECT NO. — TASK NO. — WORK UNIT ACCESSION NO.

11. TITLE (Including Security Classification)
AN EMPIRICAL ANALYSIS OF ENLISTMENT INTENTIONS AND SUBSEQUENT ENLISTMENT BEHAVIOR

12 PERSONAL AUTHOR(S)
SNYDER, Richard Paul

13 TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. Page Count
Master's Thesis	FROM TO	1990, September	99

16. SUPPLEMENTAL NOTATION
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Enlist, Propensity, Interest, YATS, Multinomial Logistics Regression

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
This research uses data from the Youth Attitude Tracking Survey and the Defense Manpower Data Center to predict interest in joining the military service (propensity) for the prime market of 17 to 21 year old high school diploma graduates that are expected to score above the fiftieth percentile on the military entrance examination. A follow-on analysis of actual conversion of propensity to enlistment action is also conducted. In predicting military interest, the independent variables were restricted to those that have data available on a regional level. This will enable military recruiting commands to develop regional estimates of propensity. Multinomial logistic regression was used to estimate the interest prediction equations for propulation groups by race and gender. Interest categorization was possible with reasonable accuracy using local unemployment level, parent's education and the regional "go to college" rate as the independent variables. Conversion of military interest to enlistment action does appear to vary by interest level. Follow-on research and recommendations are provided.

20 DISTRIBUTION/AVAILABLTIY OF ABSTRACT X UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC	1a. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Linda Gorman	22b. TELEPHONE (Include Area Code) (408)646-2767 — 22c. OFFICE SYMBOL AS/Gr

DD Form 1473, JUN 86 — Previous editions are obselete. — SECURITY CLASSIFICATION OF THIS PAGE Unclassified
S/N 0102-LF-014-6603

Approved for public release; distribution is unlimited

An Empirical Analysis of Enlistment Intentions and Subsequent Enlistment Behavior

Richard Paul Snyder
Lieutenant, United States Navy
B.S., Tulane University, 1983

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1990

Author: Richard Paul Snyder

Approved By: George W. Thomas
George W. Thomas, Thesis Co-Advisor

Linda Gorman
Linda Gorman, Thesis Co-Advisor

Dan C Boger
Dan C. Boger, Second Reader

P. Purdue
Peter Purdue, Chairman
Department of Operations Research

ABSTRACT

This research uses data from the Youth Attitude Tracking Survey and the Defense Manpower Data Center to predict interest in joining the military service (propensity) for the prime market of 17 to 21 year old high school diploma graduates that are expected to score above the fiftieth percentile on the military entrance examination. A follow-on analysis of actual conversion of propensity to enlistment action is also conducted. In predicting military interest, the independent variables were restricted to those that have data available on a regional level. This will enable military recruiting commands to develop regional estimates of propensity. Multinomial logistic regression was used to estimate the interest prediction equations for population groups by race and gender. Interest categorization was possible with reasonable accuracy using local unemployment level, parent's education and the regional 'go to college' rate as the independent variables. Conversion of military interest to enlistment action does appear to vary by interest level. Follow-on research and recommendations are provided.

TABLE OF CONTENTS

I. INTRODUCTION

Demographers have forecast a continuing decline in the number of eligible military applicants available until the mid 1990's (Office of Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics), 1978, p.183). This translates into increasingly difficult recruiting for the military. Key to the efficient allocation of recruiting efforts is the existence of an accurate measure of recruit market potential; this allows for the optimum placement of recruiting resources.

Since the inception of the military's all volunteer force, the recruiting of qualified personnel has been an important issue for the armed forces. The demanding task of seeking out qualified recruits has provided a fertile area for manpower research. Current world affairs do not lessen the importance of research in the armed forces recruiting field. If the required number of personnel is reduced, the services will still need to find the highest quality recruits available to fill those reduced billets, and the armed services will need to improve the efficiency of their recruiting efforts as their budgets dwindle. Using multinomial logistic regression, this thesis will develop a means to estimate the potential size of the qualified and interested population of potential recruits for a specific geographic market area. If the armed

forces can determine 'a priori' the level of interest in the military for a given geographic area, resources can be more effectively allocated to those areas.

Past research has concentrated on analyzing enlistment interest or propensity under the assumption that propensity is a predictor of enlistment probability (Orvis, 1982). However, determining enlistment propensity is just one aspect of modeling enlistment. Conversion of this propensity into actual enlistment is an important indicator of propensity prediction accuracy. Conversion rates and interest in joining the military will vary among different sub-groups of possible recruits. It is only when propensity and conversion rates are analyzed that a more accurate measure of market potential can be determined.

The primary objective of this research is to analyze both propensity for military enlistment and the subsequent conversion of that interest to provide the needed measure of market potential. An individual is defined to have converted interest to action if they take an enlistment entrance test, actually enlist or enter the delayed entry program for later enlistment. This analysis is done by developing a quantitative model that is used to predict the expected level of interest in enlisting in the military. Actual enlistment behavior is then analyzed by level of interest. The prediction of interest uses demographic data available on a

regional level. There are several subsidiary questions that arise in this analysis.

- What are the factors and individual characteristics that explain the enlistment propensity of non-prior service youths with varying levels of expressed interest?
- Is the expressed level of interest in enlistment an accurate indicator of future enlistment behavior?
- If propensity for military service can be determined, what is the likelihood, given a certain propensity, for actual service enlistment?
- Can a prediction of propensity be combined with the appropriate conversion rates on a regional level thus providing a measure of the number of expected recruits?

In analyzing factors and characteristics that may predict enlistment propensity, this research will consider only those explanatory variables for which data are readily available on a regional level. Thus, recruiting commands can use these results to estimate market potential for recruiting regions. Although this may degrade the predictive ability for enlistment interest relative to models using more specialized data, it will ensure that a usable model is developed.

This research is limited to the **prime market**; non-prior service youths between the ages of 17 and 21, with high school diplomas, and expected Armed Forces Qualification Test (AFQT) scores above the fiftieth percentile; mental groups I through IIIA (MG I-IIIA). All the services administer the AFQT as a required entrance examination to try to measure a recruit's trainability. Raw scores are converted to percentiles. The

scores are then converted to mental groups I through V. Table 1.1 shows those mental groups and the corresponding AFQT percentiles.

TABLE 1.1
AFQT AND MENTAL GROUP CLASSIFICATION

AFQT Percentile	Mental Group
93-99	I
65-92	II
50-64	IIIa
31-49	IIIb
21-30	IVa
16-20	IVb
10-15	IVc
1-9	V

Those without high school diplomas and those in the lower mental groups (MG IIIB-V) are not difficult to recruit. Historically, the quantity of these people supplied exceeds the quantity demanded by the services. (The number of non-high school graduates and minimum acceptable mental aptitude levels are mandated by Congress to the services.)

The data used to develop this model are from the Youth Attitude Tracking Survey (YATS) for the years 1984 to 1989 and from Defense Manpower Data Center (DMDC) personnel files. The YATS contains over 300 variables with answers to questions about education, employment, family background, military awareness and attitudes towards military service. The YATS data are collected in annual telephone surveys of a sample of

continental United States residents who are between the ages of 16 and 24, without prior military service and who have completed less then two years of college. The respondents are grouped into four basic market segments; young males (16-22 years old), older males (22-24 years old), young females (16-22 years old) and older females (22-24 years old).

These data have been matched with personnel files from DMDC if a YATS respondent actually enlisted, entered into the delayed entry program, or completed an entrance examination. The information from DMDC includes the date of enlistment or entrance testing, demographic information and aptitude test (AFQT) scores. There is a possible source of bias in the matched data set, in that file matching is possible only when the individual in the YATS provided a social security number. Roughly sixty percent of the respondents provided a social security number. The respondents that did not provide a social security number may not be representative of the sample as a whole. They are likely to be younger or without employment history. In future research, this may become less of a problem as tax laws currently require a social security number for all claimed dependents over the age of two. There will continue, however, to be those that refuse to provide this information.

In summary, this research uses YATS data to provide an equation to predict the level of interest in joining the military for those in the 'prime' recruit market. The next

part of the research is an analysis of actual enlistment behavior or conversion for that level of interest. Combining both interest and conversion should provide a more accurate estimate of the potential market for recruits. This research will not explore the size of the qualified market available (QMA). (See Peterson, 1990 for a discussion of QMA.) The purpose in this analysis is to provide an estimate of the interested or potential market available also called the qualified military interested (QMI). When both QMA and QMI are considered for an area, the military recruiting commands will have a more accurate estimate of the number of interested and qualified potential recruits and can better allocate their recruiting resources.

II. LITERATURE REVIEW

There has been a significant amount of previous research directed towards propensity analysis. Why does this thesis provide valuable new information? There are three primary unique aspects of this research that make it significantly different from past research.

First, prior to propensity analysis, the sample was reduced in this research to include only those that the military is interested in recruiting, the **prime market.** Intuitively, one would expect that interest levels of those without high school diplomas and those in the lower mental aptitude group should be higher as these people have fewer employment alternatives and may view joining the military as an opportunity to better their prospects. This would skew the levels of predicted interest and preclude accurate forecasting of interest level by region for the targeted prime market.

The second unique aspect of this research is the way in which interest levels are categorized. The majority of past research categorized interest into two levels, interested and not interested. The three way classification used in this research, interested, not interested, and neutral provides more useful results. Identifying the neutral market size is important. This is the market segment towards which more recruiting efforts should be directed in the belief that those

with a strong interest are more likely to enlist regardless of recruiting effort and those that express a strong disinterest will not benefit as much from increased recruiting efforts. The conversion rates of interest to enlistment action vary significantly by interest level for all three groups. This is another justification for the three way classification.

The final different aspect of this research is the analysis of conversion rates by market segment, race and gender, and interest level. Combining predicted propensity and conversion rates provides a more accurate measure of the true market potential. Predicted interest alone does not quantify the market size unless the conversion rates by interest are also computed.

This thesis attempts to improve on past efforts in propensity analysis by identifying interest for the prime market, developing an estimating equation for three interest levels using multinomial logistic regression and regionally available independent variables, and combining interest and conversion to more accurately identify the potential market size.

Research by others has concentrated on two basic areas of propensity analysis; propensity as an independent variable in forecasting enlistments and a computation of propensity by region or demographic factors as provided in the YATS.

Initially, the majority of research concentrated on validating expressed military interest as an adequate

indicator of future enlistment. In this respect, propensity was viewed as an independent variable that could be used for forecasting the number of potential recruits. Borack summarized the econometric models used to predict enlistment and, of the 21 models tabulated, only two use military propensity as an independent variable (Borack, 1984, p.4).

A possible reason for this is that it was not until the initial work by Orvis [1982] that propensity, as expressed by respondents on the YATS, was shown to improve prediction of actual enlistments. Of the two models mentioned by Borack that used propensity as an independent variable, one was developed in 1983, after the initial Orvis research. The other model using propensity was developed prior to Orvis' verification of propensity as a predictor of enlistment behavior, and may have been less believable. Orvis wrote regarding his initial analysis:

> The results (of propensity analysis) suggest that the enlistment intention measures in the YATS surveys do a good job of discriminating the respondents' true probabilities of enlistment. (Orvis, 1982, p.v)

In 1985 Orvis and Gahart published a report that concentrated on identifying whether propensity contributed significantly more than demographic data alone in predicting enlistment. Again, the results were in favor of using propensity to predict military enlistment. They found that even if demographic factors are controlled for, expressed military interest is still a significant predictor of enlistment (Orvis and Gahart, 1985, p.16).

Several other researchers have analyzed military propensity. Both Huzar and Goldberg and Goldberg developed regional indices of military propensity. These indices were developed by comparing interest responses on the YATS for specific geographic areas to the overall national level of military interest. Huzar determined that positive propensity indices can be developed that can in turn be used to help assign recruiting goals (Huzar, 1988, p.58). Goldberg and Goldberg used the same comparative approach to develop Reserve Propensity Indices (RPI) for use in forecasting the levels of expected Army Reserve recruits by region. Sample sizes were too small for direct estimation of propensity by region so Goldberg and Goldberg grouped the data by census region and year. They were able to determine regional RPIs that did indicate varying levels of Army Reserve interest by region (Goldberg and Goldberg, 1989, p.49).

Gorman and Mehay also examined propensity to join the Army Reserve. YATS data was pooled by year for 1983-1985 and 1985-1987. The authors then calculated confidence intervals for enlistment propensity by recruiting battalion. These estimates were then used on a comparative basis to identify areas where the interest in joining the reserves is significantly higher or lower then the mean national level of propensity. The regional propensity level was then applied to the projected military qualified population to provide a measure of the propensity-weighted military available

population. The authors recommended that the propensity adjusted populations should be used to better allocate recruiters and quotas to recruiting areas (Gorman and Mehay, 1989, p.ii).

Research on military interest is not new. There are aspects of this thesis, as summarized earlier, that contribute additional information. This additional information should allow military recruiters to better estimate the size of the potential recruit market in a geographic area.

III. DATA

A. DATA DESCRIPTION

The data used in this research were obtained from two sources. The primary data are from the Youth Attitude Survey II (YATS II). When possible, the records from the YATS were matched with Defense Manpower Data Center (DMDC) files. DMDC creates a record for an individual if he takes a pre-enlistment examination, enters the delayed entry program (DEP), or enlists in any of the armed services. The complete data set which includes YATS II responses and DMDC record information, if applicable, is referred to as the matched data set.

In this analysis the data are used in two ways. First, the entire set, matched or not, is used to estimate the stated propensity to enlist in the armed forces. The second use is to calculate the rates of conversion to enlistment action for individuals with varying interest levels of enlistment propensity. Calculations of conversion can be computed only for the respondents who provided a social security number. If a YATS respondent does provide a social security number, but there is not a matched DMDC file, that person is assumed not to have taken an enlistment action.

The earliest version of YATSII dates back to 1975 and was known as YATS. The data was initially collected twice a year

in the spring and in the fall. In 1978, YATS information was combined with the information obtained from another survey known as the Reserve Component Attitude Study (RCAS) to help design military recruiting strategies. In 1981, the spring data collection was dropped and the data collection continued to be done annually in the fall. Finally, in 1983 the YATS and RCAS was combined into one survey to produce what is now known as YATSII. From 1975 to 1983, the YATS underwent numerous changes in survey questions, sampling design and weighting techniques. The survey has, however, remained fairly consistent since 1983 with the exception of a market definition change in 1986 when the older males were redefined as ages 22 to 24 rather than ages 22 to 29, and an older female group from ages 22 to 24 was added to the survey sampling frame.

The initial matching of YATS and DMDC files was done in October of 1989. Using record dates from the DMDC data, the latest enlistment actions recorded occurred in September of 1988. A second match was then requested to obtain more recent conversion actions. This second match was done in May of 1990 and provided a most recent enlistment action date of September 29, 1989. This more recent matching also allowed the use of YATSII data for the year 1989.

DMDC has provided matched data for the YATS II respondents in 1984, 1985, 1986, 1987, 1988 and 1989. This matching was done using YATSII provided social security

numbers. Approximately sixty percent of the YATSII respondents provided social security numbers. The other forty percent either did not have a social security number, did not know their number, or refused to provide that information. There are a total of 62,480 observations for all six years. Of these, approximately 3260 observations were matched to DMDC files. These observations are roughly equally spread over the years yielding about 10,414 YATS observations per year and about 544 observations that include both YATS and DMDC data per year.

The sampling design developed for the YATS II is a two stage procedure that is based on Mitofsky/Waksberg random digit dialing. The first sampling stage provides clusters of households identified by the first eight of a ten digit phone number. Once a residence is identified in a first stage cluster, the second stage is constructed using a random selection of the final two digits. Because of the necessity to penetrate specific geographic areas and market segments (younger males (16-21), older males (22-24), younger females (16-21) and older females (22-24)) at specified levels, there was a deviation from the Mitofsky/Waksberg procedure to ensure the necessary stratification of the sample. This deviation was a disproportionate allocation of samples to geographic strata defined by Military Entrance Processing Stations (MEPS) and different sampling rates for the market segments. This negates the inherent self-weighting feature of the

Mitofsky/Waksberg procedure. Another deviation was in not sampling for the final two digits with replacement. This was done to prevent the possibility of calling the same household twice. (Waksberg, 1978, pp.40-46)

Several factors were considered in developing sample sizes. The sample sizes were computed for each MEPS based on the desired market segment stratification, the mandated predictive accuracy and the cost of survey completion. Market segment stratification was generally 50 percent younger male, 10 percent older male, 30 percent younger female and 10 percent older female. Total sample size was around 10,000. A required lower bound on the predictive accuracy or response variance was also set by the services. The measure for response variance was the estimated proportion of each market segment with a positive propensity toward military service as determined from the question concerning the likelihood of serving in the military in the next few years. This estimate was required for all segments on a national level and also at the MEPS level for young males.

Finally, estimated interview costs were developed. The costs, stratification requirements, precision requirements and historic completion rate information were combined to determine final sample size by MEPS. Interview completion rates were generally 77 percent for young males, 65 percent for older males, 78 percent for young females and 70 percent for older females (Immerman, York, and Mason, 1987, p.63).

Appendix A contains a mapping of MEPS to specific service recruiting areas. It also contains the precision requirements and sample sizes for 1987 by MEPS. This sample size determination procedure was the same for the years 1984 to 1989 and the end results varied only slightly.

B. DATA REDUCTION

The primary goal of this analysis is to identify the predictive characteristics of enlistment propensity for what has been identified as the 'prime market' for recruits. This market is 17 to 21 year old people with high school diplomas that score in the upper percentiles on the military aptitude test (AFQT). This research is unique in that the sample is restricted to the primary market prior to the development of the predictive equations for enlistment propensity. The majority of recruiting efforts are directed towards the prime market; the other markets are relatively self-recruiting. In other words, high school non-graduates and lower mental group people apply for enlistment on their own in more than sufficient numbers. Additionally, those older then 21 that meet the education and aptitude requirements are likely to have established other careers and do not provide a fertile recruiting market.

The initial data set was reduced to include only those in the primary market.[1] The initial sample was over 62,000 respondents. The reduction was accomplished using information provided by the respondent. Initially, almost 1000 responses were eliminated because of invalid or missing responses.

Self-reported age was used to eliminate respondents that were 16, 22, 23, or 24 years old. There were 12,071 16 year olds, 3456 22 year olds, 3290 23 year olds, and 3178 24 year olds. When these 21,995 people were removed because of their age at survey, this reduced the sample to approximately 39,000 people.

Several other screens were required that were not as direct. Deleting all non-high school diploma graduates was a procedure that consisted of several parts. The first step was to eliminate those that were non-graduates. This was done using three questions. If the respondent answered "none" to the types of degrees he had received from schools he had attended and would not be enrolled in any school the following year, that respondent was classified as a non-graduate. This removed 3100 observations from the initial sample.

The next step was deleting high school graduates or future graduates that had or would not be receiving a high school diploma. Certificate graduates are in this category.

---

[1]Statistical Analysis Software (SAS) version 5.18 was used for calculation and data manipulation. Multinomial logistic regression was done using Categorical Modeling (CATMOD) in SAS and Multinomial Logit (MLOGIT) developed by Salford Systems for use in SAS.

A certificate holder is any respondent who received some type of high school credential other than a regular diploma. These included Adult Basic Education (ABE) certificates and General Education (GED) certificates. This was accomplished using questions 406 and 408 or 408A. (Note that question 408 was renumbered to question 408A in 1986 and remained 408A through 1989; also, Appendix B contains a listing of all the YATS questions referenced in this research.) Question 406 asked what type of high school credential the respondent possessed. If the respondent replied anything other than a high school diploma, that person was classified as a certificate holder. Also, if a respondent did not currently have any type of credential but was or will be enrolled in an ABE or GED program, as reported in question 408 or 408A, he was considered a future certificate holder and deleted from the prime market. Using these screening factors, 1730 respondents were deleted from the survey. Another deletion that was necessary was for those in vocational or apprenticeship programs. Q408/408A was once again used. If the interviewee replied that he was or would be enrolled in a vocational training program and did not yet have a high school diploma, he was deleted. There were 739 people in this category.

The next deletion was fairly involved. In the initial sample, there were 12,000 respondents that were classified as being currently in school and therefore potential high school diploma graduates. The next step was to determine which of

these could be considered future high school diploma graduates and could therefore be included in the prime market analysis. Those still in high school, but likely to graduate, are of interest to the recruiting commands because these students are the group that must be actively recruited, yet are still inexpensive to locate. The 'in school' respondents were identified using questions 407 and 408/408A. If the person replied that he would be in school and that the type of school program was a regular day high school, he was considered 'in school' and a potential high school diploma graduate.

In order to determine which of those still in high school would obtain high school diplomas, several other responses were considered. If a student had self-reported grades of B's and C's or better, he was classified as a future graduate. If a student currently had a C average he was further screened with the college entrance examination question. If he also had taken or planned to take a college entrance exam he was considered a graduate despite his C grades. Those with grades lower then C's or those with C's that had neither taken nor planned to take a college entrance examination were considered to be non-graduates. This criterion provides a conservative estimate of those in high school that are expected to graduate. This procedure allowed the classification of 10,000 of the 12,000 'in schools' as high school diploma graduates. This screening percentage is consistent with high school graduation rates (Ogle and Asalam, 1990, p.22).

A final screen was necessary to determine those that would score in the upper mental groups on the Armed Forces Qualification Test (AFQT). The services typically are interested in those recruits that can score at or above the fiftieth percentile. This is the upper mental group (I-IIIa) population that is actively recruited. Unfortunately, AFQT score is not provided in the YATS. This information was available only if DMDC was able to match a YATS record. DMDC matching occurred for five percent of the YATS records. Since the goal of this research is predicting enlistment propensity for those in the upper mental groups, a method had to be developed to determine which respondents could be expected to score above the fiftieth percentile and thus be qualified as upper mental group (I-IIIa).

Bruce Orvis and Martin Gahart predicted AFQT upper mental group probability using a two stage probit analysis and predictor variables available in YATS (Orvis and Gahart, 1989, p.8). This prediction was provided for each individual in the YATS. A major problem with their technique in the context of this research was that one of their explanatory variables was propensity to join the military. Using their formula to classify the upper mental group population would induce significant bias when propensity is then predicted for that population. A new model was needed to classify which YATS respondents would be expected to score in the upper mental groups.

A sample size of 1367 was available for analysis (limited because of the DMDC match requirement). A match was identified using the record type variable from the DMDC file. If there was a valid response for record type, then the individual was matched by DMDC and AFQT score was available for mental aptitude analysis. Redefinition of AFQT percentile was necessary to ensure proper estimation. The AFQT is a normalized and standardized test that is scored on a percentile basis. Therefore, before estimation, the standardized percentile score was converted to a raw score. The estimation was then conducted with the raw score as the dependent variable. Once the estimated equation was applied to a set of independent variables, the resulting number was then converted back to a standardized percentile.

Ordinary least squares (OLS) regression was used to estimate the fraction of respondents that would score in the upper mental groups. AFQT percentile, as provided in the DMDC match, was the dependent variable (y) in the following assumed functional form;

$$y_i=\alpha+\beta_1*x_1+\ldots+\beta_n*x_n+\mu_i.$$

The independent variables (x) were obtained from YATS responses; the α and β coefficients were then estimated. Using OLS to predict actual AFQT percentile instead of using a binary dependent variable, upper mental group or not, provided greater flexibility in defining the upper mental group population.

The initial model was developed for all the observations with a large initial number of independent variables. The initial predictor variables were chosen because of their intuitive relationship to mental aptitude and previous significance in predicting aptitude (Orvis and Gahart, 1989, pp.11-13). Twenty-one independent variables were used in the initial model. Those variables are shown in Table 3.1.

TABLE 3.1
REGRESSION VARIABLES

RACEBLK	binary variable; 1 if Black, 0 otherwise
RACEHSP	binary variable; 1 if Hispanic, 0 otherwise
MOMED	mother's education variable; 1 if less then high school graduate 2 if high school graduate 3 if some college 4 if college graduate
DADED	father's education variable; same values as MOMED
PARED	parent's education variable; the greater of MOMED or DADED
SENIOR	binary variable; 1 if high school senior, 0 otherwise
GPA	high school grades variable; 5 if mostly A's 4 if mostly A's and B's 3 if mostly B's 2 if mostly B's and C's 1 if mostly C's or lower
SKTYP	type of high school attended variable; 1 if vocational or technical 2 if commercial or business training 3 if academic or college preparatory
INCOL	binary variable; 1 if in college, 0 otherwise
INCOLEM	binary variable; 1 if in college and employed, 0 otherwise
GENDER	binary variable; 1 if male, 0 if female
XPAY	variable for the expected pay in the next full-time job (converted to an annual salary in dollars); 1 if less than 10,000 2 if 10,001 to 16,000 3 if 16,001 to 25,000 4 if 25,001 to 50,000

Courses taken in high school are also important. Table 3.2 shows the important courses defined as binary variables; 1 if that course was taken in high school, 0 if it was not taken.

TABLE 3.2
COURSES TAKEN IN HIGH SCHOOL

ALGELE	elementary algebra
ALGINT	intermediate algebra
GEO	geometry
TRIN	trigonometry
CALC	calculus
CS	computer science
PHY	physics
BM	business math

The interaction of high school courses and grade point average (GPA) is also important. Table 3.3 shows how this interaction variable was constructed.

TABLE 3.3
CONSTRUCTED REGRESSION VARIABLE

PROD	the product of CRSE times GPA where CRSE is a summation of the course variables

Several parameter estimates were significantly different from zero using a standard two-tailed t-test at an alpha level of .05. Table 3.4 shows the initial variables and the results of the significance tests.

TABLE 3.4
REGRESSION RESULTS

Variable	Estimate	t	P
INTERCEPT	-0.57	-5.14	<0.01
RACEBLK	-0.64	-14.73	<0.01
RACEHSP	-0.25	-3.92	<0.01
MOMED	0.06	2.02	0.04
DADED	0.06	1.74	0.08
PARED	-0.06	-1.29	0.20
SENIOR	0.06	1.38	0.17
GPA	-0.02	-0.59	0.56
SKTYP	0.13	6.12	<0.01
GENDER	-0.08	-1.76	0.08
XPAY	-0.05	-2.13	0.03
ALGELE	0.24	4.73	<0.01
ALGINT	0.08	1.73	0.08
GEO	0.13	2.92	<0.01
TRIN	0.24	4.39	<0.01
CALC	0.15	2.04	0.04
CS	0.13	2.99	<0.01
PHY	-0.15	-2.78	<0.01
BM	-0.03	-0.86	0.39
INCOL	0.19	4.10	<0.01
INCOLEM	-0.16	-1.27	0.20
PROD	0.01	1.46	0.14

Upon analysis of the significant variables, it was apparent that the effects of these variables may vary depending on whether the respondent was still in high school or had graduated. This effect was expected of several

variables. The course completion variables were further examined. A t-test for equality of means for the high aptitude and low aptitude groups was done on the course completion variables. With the exception of business math, the hypothesis of equal means could not be rejected at an alpha level of .05. The other variables that warranted further examination were SENIOR, INCOL and INCOLEM. Once again a t-test for equality of means for the two aptitude groups was done and the hypothesis of equal means was rejected for all three of these variables at an alpha level of .05. Based on the difference in means and expected different course variable effects on mental aptitude, despite their equality of means, another regression model was developed.

The second model allowed separate estimation of the parameters for two subsets of the original population; those still in high school at the time of the survey and those who had already graduated. This provided a sample size of 339 for those still in school and 1028 for graduates. Orvis and Gahart also divided the population in this manner in their work. Another technique would be to introduce another binary variable for in-school or not; this however, would allow only a shift in the regression line and not account for the expected varying effect of the variables. In other words, there is not a general increase or decrease in mental aptitude based on whether a respondent is in school or not. However, the effects of the significant variables may differ depending

on the school status. For these reasons, the sample was divided based on in-high-school or not and the regression coefficients were re-estimated for each group. All the initial variables were included except INCOL and INCOLEM for the still in high school group, and SENIOR for the graduated group. The results for the in-school group regressions are given in Table 3.5 and for the graduate group in Table 3.6.

TABLE 3.5
REGRESSION RESULTS FOR THE IN-SCHOOL GROUP

Variable	Estimate	t	P
INTERCEPT	-0.85	-3.86	<0.01
RACEBLK	-0.59	-7.27	<0.01
RACEHSP	-0.23	-1.97	0.05
MOMED	0.11	1.92	0.05
DADED	0.00	0.01	0.99
PARED	-0.09	-1.01	0.31
SENIOR	0.12	1.48	0.14
GPA	0.01	0.21	0.84
S[illegible]TYP	0.20	4.96	<0.01
GENDER	0.03	0.34	0.73
XPAY	-0.11	-1.94	0.05
ALGELE	0.33	3.63	<0.01
ALGINT	0.16	1.82	0.07
GEO	0.18	2.08	0.04
TRIN	0.38	3.81	<0.01
CALC	0.29	1.85	0.07
CS	0.05	0.63	0.53
PHY	-0.13	-1.22	0.22
BM	0.03	0.49	0.62
PROD	-0.00	-0.32	0.75

TABLE 3.6
REGRESSION RESULTS FOR THE GRADUATE GROUP

Variable	Estimate	t	P
INTERCEPT	-0.50	-3.82	<0.01
RACEBLK	-0.66	-12.70	<0.01
RACEHSP	-0.25	-3.28	<0.01
MOMED	0.05	1.28	0.20
DADED	0.07	1.80	0.07
PARED	-0.04	-0.77	0.44
GPA	-0.02	-0.71	0.48
SKTYP	0.10	4.13	<0.01
GENDER	-0.11	-2.13	0.03
XPAY	-0.03	-1.37	0.17
ALGELE	0.21	3.49	<0.01
ALGINT	0.06	1.08	0.28
GEO	0.11	2.08	0.04
TRIN	0.18	2.79	<0.01
CALC	0.12	1.44	0.15
CS	0.16	3.17	<0.01
PHY	-0.13	-2.20	0.03
BM	-0.05	-1.33	0.18
INCOL	0.19	3.96	<0.01
INCOLEM	-0.16	-1.25	0.21
PROD	0.02	1.80	0.07

Based on these results, a second pair of equations was developed. For the in-school group, several variables were dropped from the equation because of insignificance. The hypothesis that DADED, PARED, SENIOR, CS, PHY, PROD and BM were jointly equal to zero, using an F test, could not be

rejected at an alpha of .05. The variables GPA and GENDER were kept even though they were insignificant because of an intuitive, but uncertain, relationship between them and mental aptitude. For the graduate group MOMED, PARED, INCOLEM, XPAY, ALGINT, CALC and BM were jointly equal to zero at the 0.05 level and were dropped from the estimating equation. GENDER was again kept despite its insignificance because of the intuitive but uncertain relationship with mental aptitude.

Table 3.7 gives the results for the in-school group. Table 3.8 gives the results for the graduate group.

TABLE 3.7
REGRESSION RESULTS FOR THE IN-SCHOOL GROUP

Variable	Estimate	t	P
INTERCEPT	-0.71	-4.12	<0.01
RACEBLK	-0.59	-7.35	<0.01
RACEHSP	-0.25	-2.16	0.03
MOMED	0.05	1.31	0.19
GPA	-0.01	-0.23	0.82
SKTYP	0.18	4.67	<0.01
GENDER	-0.01	-0.15	0.88
XPAY	-0.10	-1.94	0.05
ALGELE	0.34	3.98	<0.01
ALGINT	0.12	1.51	0.13
GEO	0.17	2.29	0.02
TRIN	0.36	4.15	<0.01
CALC	0.21	1.62	0.10

TABLE 3.8
REGRESSION RESULTS FOR THE GRADUATE GROUP

Variable	Estimate	t	P
INTERCEPT	-0.50	-4.30	<0.01
RACEBLK	-0.66	-12.74	<0.01
RACEHSP	-0.26	-3.48	<0.01
DADED	0.06	2.56	0.01
GPA	-0.05	-1.64	0.10
SKTYP	0.11	4.44	<0.01
GENDER	-0.13	-2.42	0.02
ALGELE	0.19	3.26	<0.01
GEO	0.11	2.08	0.04
TRIN	0.18	2.79	<0.01
CS	0.14	2.72	0.01
PHY	-0.15	-2.48	0.01
INCOL	0.19	4.21	<0.01
PROD	0.03	3.44	<0.01

Theoretical goodness-of-fit for the model of the in-school group was good. The F-test for the hypothesis that all coefficients are equal to zero was rejected at an alpha level of .05. The R-squared value was 0.46, and the R-squared adjusted for the number of coefficients was 0.44. Analysis of the residuals showed a mean of 0.00 with a variance of 0.29. Using a Kolomogorov D statistic for a test of normality, the hypothesis of an underlying normal distribution was not rejected at the .05 alpha level.

The graduate group did not have as good a theoretical goodness-of-fit. The F-test for the hypothesis that all coefficients are equal to zero was again rejected at an alpha

level of .05. The R-squared value was 0.38, and the R-squared adjusted for the number of coefficients was also 0.38. Analysis of these residuals showed a mean of 0.00 with a variance of 0.38. Using a Kolomogorov D statistic for a test of normality, the hypothesis of an underlying normal distribution was not rejected at a .05 alpha level. Since the goal of these regression equations is to predict mental aptitude level, the most important measure for goodness-of-fit is a comparison of predicted aptitude versus actual aptitude. Using the same data from which the coefficients were estimated, the years 1985, 1986, 1987 and 1988, predicted aptitude and actual aptitude were compared. The respondent was considered upper mental group if AFQT percentile was equal to or greater then .50.

As Table 3.9 shows, the model predicted 79.0 percent correct. Since the objective is to keep only those who would score in the upper mental groups for propensity analysis, the false high aptitude prediction rate is a key number. It was 12.7 percent.

TABLE 3.9
MODEL PREDICTION RESULTS FOR THOSE IN SCHOOL

		Predicted	
		Low Mental Group	High Mental Group
Actual	Low Mental Group	118	43
	High Mental Group	28	150

Table 3.10 gives the prediction for the graduate group which was 74.8 percent correct. This is substantially the same as for the in-school group. The false high aptitude prediction rate was 15.0 percent.

TABLE 3.10
MODEL PREDICTION RESULTS FOR GRADUATES

		Predicted	
		Low Mental Group	High Mental Group
Actual	Low Mental Group	340	154
	High Mental Group	105	429

The final test for goodness-of-fit was done using out of sample data for 1984 and 1989. There were 301 observations from these two years with actual AFQT scores to verify the regression coefficients. Table 3.11 shows that this prediction was 67.4 percent correct. The false high aptitude prediction rate which was 19.6 percent.

TABLE 3.11
MODEL PREDICTION RESULTS WITH DATA NOT USED IN COEFFICIENT ESTIMATION FOR BOTH IN-SCHOOL AND GRADUATES

		Predicted	
		Low Mental Group	High Mental Group
Actual	Low Mental Group	89	59
	High Mental Group	39	114

The same out of sample data were used to check the estimation technique developed by Orvis and Gahart. The respondent was classified as upper mental group by their technique if the probability of scoring in the upper mental group was greater then 0.50. This method, provided as part of YATS, had a 67.7 percent correct prediction with a false high aptitude rate of 14.6 percent. There are three reasons why the regression model developed in this thesis is better for identifying the high aptitude high school graduates for propensity analysis despite slightly better predictions by Orvis and Gahart. First, Orvis and Gahart use military propensity as an independent variable. This would bias the propensity predictions developed later for the prime market because propensity was used to determine which of the respondents would score in the upper mental groups. Second, using the currently developed equations maintains flexibility in upper mental group definition since AFQT percentile is predicted. Finally, since this approach does not use a two stage estimating technique, it is computationally simpler.

Before applying the developed equations to the entire sample to determine the upper mental group respondents, a final change was made to reduce the false high aptitude rate. The respondent was classified as upper mental group if his predicted AFQT percentile was above 60. This is in contrast to an actual AFQT percentile of 50 required for upper mental

group classification. Table 3.12 gives the results of this change for the withheld sample.

TABLE 3.12
MODEL PREDICTION RESULTS WITH DATA NOT USED IN COEFFICIENT ESTIMATION AND PREDICTED AFQT PERCENTILE OF 60 OR HIGHER FOR UPPER MENTAL GROUP CLASSIFICATION

		Predicted	
		Low Mental Group	High Mental Group
Actual	Low Mental Group	120	28
	High Mental Group	58	95

This was a correct prediction rate of 71.4 percent. More importantly, the false high aptitude rate fell to 9.3 percent. Using the higher predicted AFQT cut score resulted in a change from 57.5 percent predicted high aptitude to 40.7 percent predicted high aptitude. This loss of 16.6 percent of the high aptitude sample is acceptable because the sample sizes are large enough for propensity analysis to warrant the loss given the dramatic drop in false high aptitude rate.

The final step applied the mental aptitude prediction equations to the remainder of the population, those without actual test scores, to determine which YATS respondents can be included in the propensity prediction analysis. There may be some sampling bias in the development of the these regression coefficients because data existed only for those high school graduates that actually took a military aptitude test. This

group, the test takers, may not be representative of the high school graduate population.

Upon completion of data screening, the final sample size was reduced to 15,152. This sample included all 17 to 21 year old high school diploma graduates or future graduates likely to score in the upper mental group by gender and ethnic categorization. Tables 3.13 and 3.14 provide a breakdown of the resulting sample sizes.

TABLE 3.13
MALE SAMPLE SIZE

Age	White	Black	Hispanic	Total
17	2993	31	119	3143
18	2296	19	98	2413
19	1748	16	93	1857
20	878	6	30	914
21	658	5	30	693
Total	8573	77	370	9020

TABLE 3.14
FEMALE SAMPLE SIZE

Females				
Age	White	Black	Hispanic	Total
17	1859	38	66	1963
18	1468	28	89	1585
19	1277	35	53	1365
20	693	14	34	687
21	511	3	18	532
Total	5754	118	260	6132

IV. METHODOLOGY

The primary analytic portion of this research uses multinomial logistic regression (also called multinomial logit regression) which is a type of limited dependent variable regression analysis. In this research, the dependent variable is categorical. The categories are expressed levels of interest in joining the military. Binary logistic regression would be applicable if there were only two categories of the dependent variable. However, when levels of interest are best categorized into three levels, multinomial logit must be used. The independent variables can be either categorical, continuous, or both.

Suppose that there are k possible categories for the dependent variable. The goal is to predict the probability, P_i, that an observation will be in category i (i=1,2,...,k). Each probability P_i is thought of as the probability of being in category i relative to the probability of being in the category chosen as the reference category, P_m. In order to express these probabilities in terms of cumulative probability distributions, the following development is necessary. Assume that,

$$\frac{P_1}{P_1+P_m}=F(\beta_1' x) , \quad (1)$$

$$\frac{P_2}{P_2+P_m}=f(\beta_2'x)\ , \qquad (2)$$

and

$$\frac{P_{m-1}}{P_{m-1}+P_m}=F(\beta_{m-1}'x)\ . \qquad (3)$$

The classification of observations is a multinomial process. Notice that for each probability or category except the reference category m, there is a unique set of estimated coefficients β_i. It can be shown that these assumptions imply that

$$\frac{P_j}{P_m}=\frac{F(\beta'_j x)}{1-F(\beta'_j x)}=G(\beta'_j)\quad (j=1,2,\ldots,m-1)\ . \qquad (4)$$

Finally, because of

$$\sum_{j=1}^{k-1}\frac{P_j}{P_m}=\frac{1-P_m}{P_m}=\frac{1}{P_m}-1\ , \qquad (5)$$

this implies that

$$P_m=[1+\sum_{j=1}^{m-1}G(\beta'_j x)]^{-1} \qquad (6)$$

and

If the distribution of the error term is assumed logistic, the probabilities can be rewritten as

$$P_j = \frac{G(\beta'_j x)}{1+\sum_{j=1}^{m-1} G(\beta'_j x)} \ . \tag{7}$$

$$P_m = \frac{1}{D}, \tag{8}$$

and

$$P_j = \frac{e^{\beta'_j x}}{D} \quad (j=1,2,\ldots,m-1) \tag{9}$$

where

$$D = 1+\sum_{j=1}^{m-1} e^{\beta'_j x}. \tag{10}$$

(Maddala, 1983, pp. 59-60)

Assume that there are n individuals from whom the coefficients can be estimated. Equations 8 and 9 provide the probabilities that an individual will fall into one of k categories, given the reference category m. If x_i is the vector of characteristics for the i^{th} observation, then the probability that the observation falls into category j is given by equation 9 and the probability of being in category m, the reference category, is given by equation 8 after substituting x_i for x in the equations. In order to estimate the 'best' values for the βs, maximum likelihood techniques must be used. The likelihood function for multinomial logistic regression is

$$L=\prod_{i=1}^{n} P_{i1}^{y_{i1}}\, P_{i2}^{y_{i2}} \ldots P_{im}^{y_{im}} \qquad (11)$$

where, $y_{ij}=1$ if the i^{th} observation is in category j

$y_{ij}=0$ otherwise.

Equation 11 is then maximized with respect to the βs using the Newton-Raphson method. Because the log of the likelihood function is globally concave, the Newton-Raphson method converges in a small finite number of iterations. Also, the negative of the inverse of the Hessian, evaluated at β, is asymptotically the variance-covariance matrix of the β estimates (Goldfeld and Quandt, 1972, pp. 5-9). There are several assumptions implicit in the use of the multinomial logit technique. First, as mentioned, the error terms are assumed to be independent and identically distributed with a logistic distribution. This assumption is necessary to allow the use of the logistic cumulative distribution function in calculating the coefficients. If the error terms are correlated, a multinomial probit model may be more appropriate (Maddala, 1983, pp.62-64).

Use of the multinomial logit technique assumes that the probabilities of categorization are dependent only on the characteristics of the individual and are not dependent on the category. A similar model, McFadden logit, is appropriate if the determination of the category is dependent on characteristics of the category as well as individual characteristics (Maddala, 1983, pp.59-61). The assumption

that the category in which an observation is classified is independent of the characteristics of that category may hold in this research. The assignment of individuals to categories should be independent of the ordering of those categories.

The final implied property inherent in the multinomial logit model is called the independence of irrelevant alternatives (Judge and others, 1985, p.770). This property assumes that the odds of a particular categorization are unaffected by the presence of additional alternatives. This implies that if two or more of the categories are close substitutes, multinomial logit may not produce legitimate results. In this research, the property should not cause a problem because there is a clear distinction between the three categories. If we allow an additional category of interest, it is reasonable to assume that the odds of classification are unaffected.

There are several goodness-of-fit measures for the multinomial logit model. A pseudo-R^2 (ρ^2) for goodness-of-fit can be calculated using the ratio of the log likelihood values for the unrestricted model, β^{UR} and a restricted model where all coefficients equal zero, β^{R}. The value of ρ^2 is defined as follows:

$$\rho^2 = 1 - \frac{L(\beta^{UR})}{L(\beta^{R})} \quad . \tag{12}$$

This value is close to zero when restricting the model does not significantly affect the log likelihood values;

$L(\beta^{UR})=L(\beta^{R})$. If removing the restriction does significantly affect the log likelihood values, $L(\beta^{UR})<<L(\beta^{R})$, then the pseudo-$R^2$ will be close to one. Another technique for measuring goodness-of-fit is a chi-square test. Using the log likelihood ratio, this measurement tests the null hypothesis that all coefficients are equal to zero or each alternative category is equally likely. (Judge and others, 1985, p.774)

The final method for measuring goodness-of-fit, and of primary importance in this research, is an analysis of the predictive ability of the coefficients. This is done by computing the percentage of individuals predicted for each category and comparing these values with the percentage in each category by chance, 1/k. In other words, this method is a comparison of predicted cell membership versus actual cell membership. (Judge and others, 1985, p.773)

V. ANALYSIS

A. DEPENDENT VARIABLE

The first step in developing the estimating equation for interest in joining the military was to determine a measure of expressed interest in military enlistment using the YATS responses. There were eleven questions common to the questionnaires from 1984 to 1989 that in one form or another ask about the respondents' military propensity. Two decisions were necessary. First, which of the eleven question's responses should be used as a measure of interest? Second, what categories of interest should be defined? The candidate questions and their range of responses are given in Table 5.1. Appendix B contains a detailed description of the questions.

Two of the questions were discarded because of possible ambiguities in their responses. Question Q683 indicated only whether joining the military had been discussed, not whether the respondent was interested in the military. Question Q692 asks about the respondents' feelings regarding serving in the military; the response may or may not be representative of his actual propensity to join the military. He may have replied that he had discussed joining the military, but it was a negative discussion that emphasized how unlikely it was that he would join the military.

TABLE 5.1
PROPENSITY VARIABLES

Question	Responses
V438JOIN	military mentioned (1) or not (0)
Q503	1-definitely 2-probably 3-probably not 4-definitely not
Q517	1 or 2-in school 3 or 4-working 5-serving in military 6-homemaker 7-other
Q522	0 to 10
Q622	yes or no
Q625	yes or no
Q628	yes or no
Q645	yes or no
Q683	yes or no
Q692	1-very favorable 2-somewhat favorable 3-neutral 4-somewhat unfavorable 5-very unfavorable
CPYATS82	1-definitely 2-probably 3-probably not 4-definitely not

Four more possibilities were also rejected after further analysis of the responses to the questions. Questions Q628, talking to a recruiter, and Q645, taking the ASVAB, had abnormally high positive interest responses in comparison with other interest questions. This is not that surprising.

Youths often take the ASVAB entrance exams for reasons unrelated to joining the military entrance. Contact with a military recruiter may occur because the recruiter seeks out an individual and not necessarily because he is interested in enlisting. The responses to questions Q622 and Q625, have you called or written for information about the military, were cross-tabulated with the responses to V438JOIN; this did not indicate that these questions were indicative of positive military propensity because, of the positive responses to question Q622, 88 percent of them did not provide a positive response to question V438JOIN. Similarly, 90 percent of the positive responses to question Q625 did not corroborate this with a positive response to V438JOIN. Questions Q628, Q645, Q622 and Q625 were excluded from further consideration as indicators of enlistment interest.

The five questions that remained for consideration, V438JOIN, Q503, Q517, Q522 and CPYATS82, were initially analyzed using principal component analysis to identify any significant combination of variables. This analysis did not provide any insight. Question Q517 was dropped when a frequency cross-tabulation indicated that of the 203 favorable responses to question Q517, 157 or 77 percent of them were assimilated by the positive responses to question V438JOIN.

In looking at the four remaining questions, it was apparent that accurate categorization into two interest categories would be difficult. Because the number of levels

of responses to the questions of interest did not readily lend themselves to a two way classification, it was decided to develop three interest categories. One was the definitely interested group. The second was the neutral or undecided group, and the final group was those that were definitely not interested.

The primary justification for the interest grouping is tied to recruiting efforts. The definitely interested group does not need as much recruiting because of its high interest level; they will probably join regardless of whether they are recruited. The definitely not interested group presents the opposite problem. They will probably not join regardless of the recruiting effort directed at them. The middle group, on the other hand, may be swayed by the attention of a recruiter. Therefore, it may be important to understand the characteristics of the neutral group.

Secondly. this three way classification was a natural way to describe the responses. The definitely interested group seemed easy to identify using responses to questions V438JOIN (yes), Q503 and CPYATS82 (definitely) and Q522 (8, 9 or 10). The definitely not interested group was determined using questions Q503 and CPYATS82 (definitely not) and Q522 (0, 1, 2 or 3). Those not classified into one of the two extreme groups were classified as neutral. Appendix C contains the if/then coding that was used in SAS to classify interest

levels as discussed. The sample sizes for the categories is given in Table 5.2.

TABLE 5.2
PROPENSITY CATEGORIZATION

INTEREST LEVEL	NUMBER(PERCENT)
DEFINITELY INTERESTED (1)	610(5.90)
NEUTRAL (2)	4607(44.52)
DEFINITELY NOT INTERESTED (3)	5130(49.58)
TOTAL	10,347(100.00)

B. INDEPENDENT VARIABLES

The initial variables used to predict interest in the military were age, unemployment levels, marital status, future educational plans, the cost of higher education, parent's education level, the number of military acquaintances or relatives in the military, high school grade point average, the type of high school attended and whether or not a college entrance examination had been taken. These variables were chosen because they historically have been significant indicators of military recruiting success, and therefore they may be significant in predicting military interest.

The goal of this research was to predict interest in the military on a regional level. Therefore, the independent variables used had to have data available on a regional level. This restriction severely limited the variables for

consideration. Also, since the goal of this research is to identify the varying interest levels from region to region, the independent variable should also have variation from one area to another. These two restrictions simplified the process of identifying the possible independent variables.

Of the initial variables considered, two were eliminated because they probably do not have significant variation across regions. Within the 17 to 21 year old group, for example, it is likely that the age distribution does not vary significantly from one area to another. The same is probably true for the marital status of this group.

Several other possible predictors of military interest were eliminated because, although they existed in YATS, they were not available on a regional level. These variables were the type of high school attended, high school grade point average, taking of a college entrance examination, and the number of military acquaintances or relatives. Although it is believed that these variables may be significant predictors of military interest, they do not have adequate proxies outside of the survey.

Given these restrictions, three variables remained for consideration as independent variables. They were unemployment level, parents' education level and whether or not the respondent was in college or planned to attend college. Each of these variables met three necessary conditions. First, the information was available from the

YATS survey. Second, reliable estimates for their values was available on a regional basis, and finally the values for the variables could be expected to vary from one region to another.

A respondent's perception of unemployment level is provided in YATS in questions Q436 and Q437 for full-time and part-time employment respectively. The correlation of this perception to actuality was possible because the actual level of unemployment in the respondent's county of residence was provided in 1985. The correlation between unemployment perceptions and actual unemployment levels was highest for Q436, the perceived difficulty of finding full-time employment. Actual unemployment rates are available on a regional basis.

Using plots of perceived difficulty of finding a job and actual unemployment rates and linear regression, a mapping of the response to question Q436 and actual unemployment levels was developed. A perceived difficulty of finding a full-time job 'not difficult at all' corresponded to a 1985 regional unemployment level of less then 7 percent. A perceived difficulty of finding a job of 'somewhat difficult' corresponded to an actual regional unemployment level of 7 to 9 percent. A perception of a 'very difficult' time of finding full time employment was indicative of an actual regional unemployment level of 9 to 11 percent. Finally, if it was perceived as 'almost impossible' to find a full-time job, this

corresponded to an actual regional unemployment level of greater then 11 percent. The variable used in the interest prediction was coded 1 to 4 with 4 corresponding to the least unemployment and 1 as the highest unemployment.

The level of parents' education was also available in YATS. There is a question asking the father's highest level of education and another asking the highest level of the mother's education. These values were years of education. The mother and father values were averaged to provide an average number of years of education for the respondent's parents. The average number was then categorized into one of four categories: 1, less then 12; 2, 12 to 15; 3, 16 to 20; 4, greater then 20. This categorization corresponds to less then a high school education, high school education but not a full college education, a college education and college graduate level education. The average adult education level for a region is readily available data that can serve as a proxy for the level of parent's education.

The final independent variable was whether or not the respondent was in college or expressed a strong intention of attending college. Because of the large number of people that when asked if they would attend college replied yes (92 percent), a more involved screen was used. If a respondent replied that he intended to attend college, and that he had taken a college entrance test, and that he believed that college costs would exceed 1000 dollars per year (a proxy for

the respondent's research into college attendance), that person was identified as a future college student. If the respondent replied that he was currently in college, he was also included in this group. This screen identified 67 percent of the respondents as likely future college students. The variable was coded 1 if the respondent was likely to attend college and 0 otherwise.

The regional proxy for the 'go to college' variable is provided by the Department of Education (Hershberger, 1990). Unfortunately, further information about this data was unavailable, and Major Hershberger expressed concern about the validity of these data. Since this variable does contribute to the prediction of interest it was included although it may not be available at the regional level.

C. PROPENSITY ANALYSIS

The major thrust of this research was to develop an estimating equation for interest in joining the military for those in the prime market. The method used was multinomial logistic regression with a dependent variable consisting of three categories: definitely interested, definitely not interested and neutral interest. The reference dependent category was the neutral interest group. The available independent variables were unemployment level, parents' average education level and whether or not the respondent intended to attend college. These variables meet the criteria of regional data availability and YATS availability.

In order to allow for more flexibility in coefficient estimation, the sample was subdivided by race and gender. This created six sub-populations; white (actually white, oriental, etc. but it will be referred to as white) males, black males, hispanic males, white (white, oriental, etc.) females, black females and hispanic females. Another technique to account for race and gender would be to introduce dummy variables for gender and race. This only allows a shift in the estimated interest level function, however, and does not allow for varying coefficient effects by race and gender. Sample sizes for maximum likelihood estimation for the coefficients for each group are presented in Table 5.3.

TABLE 5.3
SAMPLE SIZE FOR MULTINOMIAL
LOGISTIC REGRESSION

Population	Sample Size
White Males	8496
Black Males	76
Hispanic Males	367
White Females	5688
Black Females	115
Hispanic Females	258

Table 5.4 provides the coefficient estimates, standard error of the estimate, the Chi-square statistic for significance at an alpha level of .05 and the probability that the coefficient is not significantly from zero. There are two values for each independent variable. The first value is used for estimating the probability of being in the definitely

interested category, and the second coefficient value is for calculating the probability of being in the definitely not interested category. (see equations (9) and (10) in the methodology chapter) The neutral probability can be calculated using the fact that the three probabilities must add to one.

TABLE 5.4
WHITE MALES

Variable	Estimate	Standard Error	Chi-Square	Probability
Intercept	-0.97	0.17	30.66	<0.01
	-0.67	0.10	40.70	<0.01
Unemployment	-0.13	0.05	8.11	<0.01
	0.09	0.03	12.01	<0.01
Parent's Education	-0.09	0.04	4.82	0.02
	0.04	0.02	3.79	0.05
Go to College	-0.41	0.08	22.91	<0.01
	0.07	0.05	1.88	0.17

With the exception of the going to college coefficient for the not interested category, all the parameter estimates are significantly different from zero at an alpha level of 0.05 percent. Using analysis of variance and likelihood ratios, Table 5.5 gives ANOVA results for the white males.

TABLE 5.5
ANALYSIS OF VARIANCE FOR THE WHITE MALE GROUP

Source	Degrees of Freedom	Chi-Square	Probability
Intercept	2	57.32	<0.01
Unemployment	2	26.51	<0.01
Parent's Education	2	11.33	<0.01
Go to College	2	30.0	<0.01
Residual Variance	56	75.35	0.04

The analysis of variance table emphasizes the fact that the effects of the explanatory variables are significant. However, the hypothesis that the residual unexplained variance is not significantly different from zero is rejected at an alpha level of 0.05 indicating that there is residual variance that could be explained.

The real indicator of estimating success is the number of correct categorizations using the computed coefficients. Because multinomial logit provides probabilities of an individual being in one of the three interest categories, two tables are necessary to summarize the predictive accuracy of the equations. The probabilities for individuals with identical characteristics were multiplied times the number of individuals with those characteristics to determine the expected number in each interest category. This expected number was then compared to the actual number in each category. Suppose, for example, for a set of characteristics

there were three people that were interested, four people that were neutral and three people that were not interested. Using the predicted probabilities for each category and multiplying times ten it is expected that two people should be interested, four people should be neutral and four people should be not interested, that prediction would be 66.6 percent (2/3) correct for the interested group, 100.0 percent (4/4) correct for the neutral group and 75.0 percent (3/4) correct for the not interested group. Basically, if there are X people in an interest category and the equation predicted Y people with that interest level, the lesser of the two values was used as the number of correct predictions. This was the technique used to compute Table 5.6 for white males. Interest level predictions were also aggregated across all sets of predictive characteristics. This does not account for correct predictions by characteristic sets as in the first technique, but it does provide an overall indicator of category prediction success. This overall prediction was used to yield Table 5.7 for white males. The same prediction rate computations were used for the other race and gender groups.

TABLE 5.6
WHITE MALE PREDICTIONS

Category	Actual Count	Correct Predictions	Percent Correct
Interested	664	607	91.4
Not Interested	3505	3410	97.3
Neutral	4327	4238	97.9
Total	8496	8255	97.2

TABLE 5.7
CATEGORIZATION FOR WHITE MALES

Category	Predicted (percent)	Actual (percent)
Interested	659(7.7)	664(7.8)
Not Interested	3508(41.3)	3505(41.2)
Neutral	4329(50.9)	4327(50.9)
Total	8496(100.0)	8496(100.0)

The black male group had the smallest sample size and the effect on coefficient significance was apparent. The estimates for this group are given in Table 5.8.

TABLE 5.8
BLACK MALES

Variable	Estimate	Standard Error	Chi-Square	Probability
Intercept	-1.48	1.90	0.61	0.43
	-1.43	1.17	1.49	0.22
Unemployment	-0.03	0.46	0.00	>0.95
	0.09	0.27	0.11	0.74
Parent's Education	-0.02	0.36	0.00	>0.95
	0.41	0.23	3.12	0.08
Go to College	0.15	1.20	0.02	0.90
	0.17	0.70	0.06	0.80

The low significance of the individual variables makes it difficult to comment on the individual effects of each variable. However, the analysis of variance indicates that the model does explain a significant portion of the variation in military interest as indicated in Table 5.9 by the inability to reject that the residual variance is significantly different from zero at an alpha level of 0.05.

TABLE 5.9
ANALYSIS OF VARIANCE FOR THE BLACK MALE GROUP

Source	Degrees of Freedom	Chi-Square	Probability
Intercept	2	1.68	0.43
Unemployment	2	0.14	0.93
Parent's Education	2	3.47	0.18
Go to College	2	0.06	0.97
Residual Variance	36	38.45	0.36

The predictive ability of the equation for black males also supports the hypothesis that the model is able to predict military interest. Tables 5.10 and 5.11 summarize the results of that prediction.

TABLE 5.10
BLACK MALE PREDICTIONS

Category	Actual Count	Correct Predictions	Percent Correct
Interested	7	3	42.8
Not Interested	38	31	81.6
Neutral	31	26	83.9
Total	76	60	78.9

TABLE 5.11
CATEGORIZATION FOR BLACK MALES

Category	Predicted (percent)	Actual (percent)
Interested	5(6.6)	7(9.2)
Not Interested	38(50.0)	38(50.0)
Neutral	33(43.4)	31(40.8)
Total	76(100.0)	76(100.0)

Although this model had some difficulty in predicting correctly for those that are in the interested group, the estimation of category proportions in Table 5.11 was much closer to the actual categorization that occurred. A larger sample size for estimation would probably produce a better correct classification rate.

The coefficient estimation results for the hispanic male group are in Table 5.12. Again, the smaller sample size makes it difficult to comment on the individual effects of each

variable. However, the analysis of variance does indicate that the unemployment variable is significant in explaining military interest for hispanic males. The likelihood ratio test for the significance of the residual variance does indicate that there was more variance that could be explained. The results of analysis of variance are presented in Table 5.13.

TABLE 5.12
HISPANIC MALES

Variable	Estimate	Standard Error	Chi-Square	Probability
Intercept	0.94	0.72	1.69	0.19
	0.10	0.50	0.04	0.83
Unemployment	-0.90	0.22	17.19	<0.01
	-0.04	0.13	0.09	0.76
Parent's Education	0.032	0.18	0.03	0.86
	-0.09	0.10	0.82	0.36
Go to College	-0.59	0.41	2.04	0.15
	-0.25	0.24	1.02	0.31

TABLE 5.13
ANALYSIS OF VARIANCE FOR THE HISPANIC MALE GROUP

Source	Degrees of Freedom	Chi-Square	Probability
Intercept	2	1.73	0.42
Unemployment	2	17.83	<0.01
Parent's Education	2	0.97	0.61
Go to College	2	2.49	0.29
Residual Variance	54	74.42	0.03

The predictive results for hispanic males are summarized in Tables 5.14 and 5.15.

TABLE 5.14
HISPANIC MALE PREDICTIONS

Category	Actual Count	Correct Predictions	Percent Correct
Interested	32	20	62.5
Not Interested	133	126	94.7
Neutral	202	186	92.1
Total	367	332	90.5

TABLE 5.15
CATEGORIZATION FOR HISPANIC MALES

Category	Predicted (percent)	Actual (percent)
Interested	30(8.2)	32(8.7)
Not Interested	134(36.5)	133(36.2)
Neutral	203(55.3)	202(55.0)
Total	367(100.0)	367(100.0)

The model for hispanic males was able to categorize the interest levels well and, like the black males, suffered from a small sample in the interested group. This was evidenced in the small fraction of correct predictions in that group.

The estimations for the females followed much the same pattern as for the males. More accurate coefficients were estimated for the white females due to a larger sample. The coefficients and their associated statistics are presented in Table 5.16.

TABLE 5.16
WHITE FEMALES

Variable	Estimate	Standard Error	Chi-Square	Probability
Intercept	-2.01	0.36	31.38	<0.01
	0.43	0.12	12.29	<0.01
Unemployment	-0.18	0.10	3.40	0.06
	0.07	0.03	4.55	0.03
Parent's Education	0.00	0.09	0.00	>0.95
	0.05	0.03	3.15	0.08
Go to College	-0.09	0.19	0.24	0.62
	0.01	0.06	0.02	0.90

Although these estimates were not as precise as for the white males, they were accurate enough for coefficient comparison in the conclusions portion of this research. Analysis of variance indicates the significance of unemployment in predicting interest and the residual variance is not significantly different form zero at an alpha level of 0.05. Table 5.17 presents the results.

TABLE 5.17
ANALYSIS OF VARIANCE FOR THE WHITE FEMALE GROUP

Source	Degrees of Freedom	Chi-Square	Probability
Intercept	2	55.45	<0.01
Unemployment	2	10.22	<0.01
Parent's Education	2	3.31	0.19
Go to College	2	0.30	0.86
Residual Variance	56	44.59	0.86

The predicted categorization results are summarized in Tables 5.18 and 5.19.

TABLE 5.18
WHITE FEMALE PREDICTIONS

Category	Actual Count	Correct Predictions	Percent Correct
Interested	132	111	84.1
Not Interested	380	373	98.3
Neutral	1755	1692	96.4
Total	5688	5540	97.4

TABLE 5.19
CATEGORIZATION FOR WHITE FEMALES

Category	Predicted (percent)	Actual (percent)
Interested	132(2.3)	132(2.3)
Not Interested	3801(66.8)	3801(66.8)
Neutral	1755(30.8)	1755(30.8)
Total	5688(100.0)	5688(100.0)

As with the males, coefficient estimation for the blacks and hispanics was less precise. These coefficients are in Tables 5.20 and 5.21 for the female blacks and hispanics.

TABLE 5.20
BLACK FEMALES

Variable	Estimate	Standard Error	Chi-Square	Probability
Intercept	2.59	1.60	2.61	0.111
	0.80	1.01	0.63	0.43
Unemployment	-0.44	0.42	1.10	0.29
	-0.17	0.23	0.53	0.47
Parent's Education	-0.71	0.41	2.94	0.09
	0.12	0.20	0.38	0.54
Go to College	-1.84	0.90	4.23	0.04
	-0.31	0.68	0.21	0.64

TABLE 5.21
HISPANIC FEMALES

Variable	Estimate	Standard Error	Chi-Square	Probability
Intercept	-0.38	1.22	0.10	0.76
	-0.53	0.61	0.76	0.38
Unemployment	-0.23	0.34	0.44	0.51
	0.21	0.15	1.90	0.17
Parent's Education	-0.05	0.30	0.03	0.86
	0.23	0.13	3.25	0.07
Go to College	-1.38	0.63	4.84	0.03
	-0.11	0.31	0.13	0.72

Analysis of variance for these two groups indicates that the residual unexplained variance for both groups is not significantly different from zero at an alpha level of 0.05. Tables 5.22 and 5.23 summarize these results.

TABLE 5.22
ANALYSIS OF VARIANCE FOR THE BLACK FEMALE GROUP

Source	Degrees of Freedom	Chi-Square	Probability
Intercept	2	2.63	0.27
Unemployment	2	1.25	0.54
Parent's Education	2	4.25	0.12
Go to College	2	4.57	0.10
Residual Variance	40	41.74	0.39

TABLE 5.23
ANALYSIS OF VARIANCE FOR THE HISPANIC FEMALE GROUP

Source	Degrees of Freedom	Chi-Square	Probability
Intercept	2	0.76	0.68
Unemployment	2	3.05	0.22
Parent's Education	2	3.70	0.16
Go to College	2	4.95	0.08
Residual Variance	50	68.05	0.05

Finally the prediction results for the black and hispanic females are provided in Tables 5.24 through 5.27.

TABLE 5.24
BLACK FEMALE PREDICTIONS

Category	Actual Count	Correct Predictions	Percent Correct
Interested	9	4	44.4
Not Interested	63	58	92.1
Neutral	43	36	83.7
Total	115	98	85.2

TABLE 5.25
CATEGORIZATION FOR BLACK FEMALES

Category	Predicted (percent)	Actual (percent)
Interested	8(6.9)	9(7.8)
Not Interested	66(57.4)	63(54.8)
Neutral	41(35.6)	43(37.4)
Total	115(100.0)	115(100.0)

TABLE 5.26
HISPANIC FEMALE PREDICTIONS

Category	Actual Count	Correct Predictions	Percent Correct
Interested	13	8	61.5
Not Interested	153	122	79.7
Neutral	92	73	79.3
Total	258	203	78.7

TABLE 5.27
CATEGORIZATION FOR HISPANIC FEMALES

Category	Predicted (percent)	Actual (percent)
Interested	12(4.6)	13(5.0)
Not Interested	149(57.7)	153(59.3)
Neutral	97(37.6)	92(35.6)
Total	258(100.0)	258(100.0)

Despite limited sample sizes for hispanics and blacks, the lowest level of correct categorizations was 78.7 percent for hispanic females. White males were correctly categorized 97.2 percent of the time. These results may be optimistic. There was no out of sample data available to further verify the predictive ability of the equations.

D. CONVERSION RATES

The ratios of enlistment interest to actual enlistment action were computed using the information available in the matched data set. If a YATS respondent provided a social security number, that person was considered a potential new recruit. If a matched military record was available for that person, he was assumed to have taken an enlistment action. The enlistment action was either testing for enlistment, actual enlistment or entry into the delayed entry program.

Of the 15,152 people in the YATS data set in the primary market, 10,347 people or 68 percent provided social security numbers. This is the sample size for conversion analysis. Of

these, 649, or 6 percent, actually completed some enlistment action. Of these, 441, or 68 percent, of the enlistment actions were test for enlistment, 183, or 28 percent, of the actions were actual enlistments and the remaining 25, or 4 percent, were entries into the delayed entry program.

The numbers desired are conversion rates for different expressed interest levels broken down by market segment. The conversion rate, or estimate of p, was calculated as

$$p=\frac{N_{enlist\ action}}{N_{with\ social\ security\ number}}.$$

A confidence interval for this computed proportion was also calculated. If each observation is considered to come from a Bernoulli distribution which is distributed BINOMIAL (1,p), the maximum likelihood estimate for p is

$$\hat{p}=\sum\frac{X_i}{n}.$$

There is no pivotal quantity to use for p to develop a confidence interval; however, the central limit theorem can be used to show that

$$\frac{\hat{p}-p}{\sqrt{p(1-p)/n}}\rightarrow Z\sim NORMAL(0,1).$$

Using this, it can be shown that an approximate confidence interval for p is

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} \quad \text{for } n\hat{p}>5 \;,\; n(1-\hat{p})>5$$

(Bain and Engelhart, 1987, p.345). Point estimates and .05 percent alpha level confidence intervals for conversion rates were computed by race, gender and interest level and are reported in the following tables.

Conversion rates for all races and both genders by interest are given in Table 5.28. Tables 5.29 and 5.30 contain conversion rates by gender. It should be noted that the sample size of white males was almost 6000. This means that any aggregated table that includes white males will be most representative of the conversion actions of white males. Conversion rates when the sample is segmented by race are given in Tables 5.31, 5.32, and 5.33. Because of the reduced sample sizes for blacks and hispanics, a conversion rate was computed for all respondents of the race regardless of interest level. Blacks had an overall conversion rate of 6.50 percent with a confidence interval of [2.14,10.86] percent. Hispanics had an overall conversion rate of 7.88 percent and a confidence interval of [5.26,10.50] percent.

TABLE 5.28
CONVERSION RATES FOR ALL RACES, BOTH GENDERS

N=10347		
	Point Estimate	Confidence Interval
High Interest n=610	21.15%	[17.91,24.39]
Neutral Interest n=4607	7.01%	[6.27,7.75]
Disinterest n=5130	3.84%	[3.31,4.37]

TABLE 5.29
CONVERSION RATES MALES, ALL RACES

N=6176		
	Point Estimate	Confidence Interval
High Interest n=494	22.67%	[18.98,26.36]
Neutral Interest n=3239	8.46%	[7.50,9.42]
Disinterest n=2443	5.94%	[5.00,6.88]

TABLE 5.30
CONVERSION RATES FEMALES, ALL RACES

N=4171		
	Point Estimate	Confidence Interval
High Interest n=116	14.66%	[8.26,21.06]
Neutral Interest n=1368	3.58%	[2.60,4.56]
Disinterest n=2687	1.94%	[1.42,2.46]

TABLE 5.31
CONVERSION RATES WHITE, BOTH GENDERS

N=9817		
	Point Estimate	Confidence Interval
High Interest n=569	21.09%	[17.74,24.44]
Neutral Interest n=1368	6.99%	[6.23,7.75]
Disinterest n=2687	1.92%	[1.39,2.45]

TABLE 5.32
CONVERSION RATES BLACK, BOTH GENDERS

N=123		
	Point Estimate	Confidence Interval
High Interest n=11	9.09%	***
Neutral Interest n=47	4.26%	***
Disinterest n=65	7.69%	***

TABLE 5.33
CONVERSION RATES HISPANIC, BOTH GENDERS

N=406		
	Point Estimate	Confidence Interval
High Interest n=30	26.67%	***
Neutral Interest n=197	8.12%	***
Disinterest n=179	4.47%	***

*** The sample sizes were insufficient to develop meaningful confidence intervals. The point estimates are subject to wide variance.

With further market segmentation, meaningful estimates for conversion rate by interest level were available only for white males and females because of sample size. These are shown in Tables 5.34 and 5.35.

TABLE 5.34
CONVERSION RATES WHITE MALES

N=5896		
	Point Estimate	Confidence Interval
High Interest n=469	22.17%	[18.41,25.93]
Neutral Interest n=3092	8.38%	[7.41,9.35]
Disinterest n=2335	5.78%	[4.83,6.73]

TABLE 5.35
CONVERSION RATES WHITE FEMALES

N=3921		
	Point Estimate	Confidence Interval
High Interest n=100	16.00%	[8.82,23.18]
Neutral Interest n=1270	3.62%	[2.59,4.65]
Disinterest n=2551	1.92%	[1.39,2.45]

Conversion rates for the markets of interest were developed from those that provided a social security number in YATS and subsequently completed some type of enlistment action. There are two sources of bias in this method. First, since we could analyze only those that provided social security numbers, there is an implied assumption that they are representative of the entire population. This may be an

unrealistic assumption. If respondents did not have a social security number, it may be that they are younger or without work experience. If the respondent refused to provide his social security number, he may covet privacy and be a less likely military candidate. The net effect of these two aspects of the social security number induced bias may or may not be zero. In any case, this bias could not be removed and the results must be tempered by the possibility of bias.

The second source of possible bias in rate estimation is the effect of time. YATS survey were available for 1984 to 1989. The match with DMDC records was done in the Spring of 1990. This implies that the YATS respondents from the earlier years had more time to complete an enlistment action. This would imply that the conversion rates should be highest for the early years and decrease until reaching a minimum for the YATS respondents from 1989. This did occur. Overall conversion rates by year are in Table 5.36.

TABLE 5.36
OVERALL CONVERSION RATES BY YEAR

Year	Rate
1984	1.41%***
1985	9.01%
1986	9.69%
1987	7.68%
1988	4.20%
1889	4.88%

***Some anomaly may exist in the 1984 match. It should be used with caution.

This time effect will also cause some bias in the conversion results. The YATS respondents in years 1987, 1988 and 1989 have apparently not had enough time to complete contemplated enlistment actions. Therefore, the conversion rates computed, which combine data from all six years, are probably lower than can actually be expected. Removing these years from the analysis would reduce sample size too severely. Thus, the biases induced by their presence must be acknowledged and accepted in this research.

VI. CONCLUSIONS

A. MILITARY PROPENSITY

The primary goal of this research was to determine the feasibility of developing equations that could predict the military interest for the prime market consisting of 17 to 21 year old high school diploma graduates that could be expected to score in the upper percentiles on the Armed Forces Qualification Test. Several results from the analysis are significant. First is the interest categorization by race and gender. Second are the estimated coefficient values and the differences in coefficients by sub-population. The final significant point is the predicted categorization of military interest.

Actual interest categorization by race and gender does appear to differ as summarized by the percentage figures in Table 6.1. Across races, the males seem more interested in joining the military. Also, hispanics and black seem more interested in the military then whites. Because of the limited sample sizes for hispanics and blacks, only a white male and white female comparison can be made with statistical significance.

TABLE 6.1
MILITARY INTEREST C\TEGORIZATION

Category	White Males	White Females	Black Males	Black Females	Hispanic Males	Hispanic Females
Interested	7.8%	2.3%	9.2%	7.8%	8.7%	5.0%
Not Interested	41.2%	66.8%	50.0%	54.8%	36.2%	59.3%
Neutral	50.9%	30.8%	40.8&	37.4%	55.0%	35.6%
N	8496	5688	76	115	367	258

If the sample size for the proportion in a category is greater than 50, the following test statistic can be computed;

$$\hat{z}=\frac{|\hat{p}_1-\hat{p}_2|}{\sqrt{\hat{p}(1-\hat{p})\,[(\frac{1}{n_1})+(\frac{1}{n_2})]}}$$

where

$$\hat{p}_1=\frac{x_1}{n_1}$$
$$\hat{p}_2=\frac{x_2}{n_2}$$
$$\hat{p}=\frac{(x_1+x_2)}{(n_1+n_2)}$$

and

$$\hat{z}\sim N(0,1)$$

(Bain and Engelhart, 1987, p.383). Using this test statistic, the hypothesis that interest categorization for white males and white females was equal was tested. The hypothesis was rejected for all three interest categories implying that there is a difference in interest by gender for whites. It is

likely that interest would also vary by gender for the other races, but sample size precludes a statistical test.

Sample size also limits the statistical testing of coefficient values. Using an alpha level of .05, the coefficients that were significant were analyzed. As expected, higher unemployment increased the probability of being interested in the military as did a lower education level for the parents. Also, if the respondent was likely to attend college, he was less likely to be interested in the military.

A comparison of coefficients by gender and race was possible only for the unemployment coefficient for white males and females. In this case, the values were not significantly different. The intercept terms were significantly different indicating that the estimating equations may differ only by a shift parameter that could be introduced using a binary variable for gender. This option could be more fully explored with larger sample sizes that would allow a more complete coefficient comparison.

Finally, the analysis of predicted categorization indicates that the model will work. These preliminary results show that all sub-populations by race and gender can be accurately categorized by level of interest in the military. The largest discrepancy was for black females where the model predicted 57.4 percent would not be interested in the military when in fact only 54.8 percent were not interested.

The information provided on military interest prediction has use for the recruiting commands. It is apparent that interest in the military will vary by geographic region as parents' education, unemployment levels, and the proportion of people interested in college varies by geographic area. Once these variables have been determined for a region, the estimating equation for each population by gender and race can be applied to determine the expected levels of interest in joining the military. This information can then be used to help determine recruiting resource allocation.

B. CONVERSION RATES

Conversion of military interest to actual enlistment action (testing, enlisting, or entering the delayed entry program) appears to differ by expressed interest in the military and less so by race and gender. Table 6.2 summarizes the results by interest level. These percentages were significantly different by interest level using the proportion hypothesis test given in section V.D. above.

TABLE 6.2
CONVERSION RATES FOR THE ENTIRE SAMPLE

	Percent Converting
High Interest	21.2
Neutral	7.0
Not Interested	3.8

Table 6.3 gives a comparison of conversion rates for white males and females. These conversion rates did vary significantly by gender for whites for the neutral and not interested categories. The difference in conversion rate for the interested group was not significantly different. Once again, the proportion hypothesis test was used to determine significant differences.

TABLE 6.3
CONVERSION RATE COMPARISON

	White Males (percent)	White Females (percent)
High Interest	22.2	16.0
Neutral	8.4	3.6
Not Interested	5.8	1.9

Of interest in conversion rate analysis is the fact that the conversion rate for the not interested group for all races and both genders was almost one seventh the rate of the interested group. However, in terms of absolute numbers, 129 of the interested group performed an enlistment action while 197 of the not interested group completed some enlistment action. This emphasizes the fact that the not interested group should not be ignored in recruiting efforts. Past studies have also pointed this out (Orvis and Gahart, 1985, p.19). The low conversion rate of the not interested group is offset by the large numbers in that group.

As stated earlier, the neutral interest group is important. This is the group which may provide the greatest increase in recruits from an increase in recruiting effort. This group, like the not interested group, does include a large proportion of the population. Considering the higher conversion rates of the neutral interest people in relation to the not interested group, the large number of people in the neutral category and the expected benefits from an increased recruiting effort re-emphasizes the importance of identifying and understanding the characteristics of the neutral interest category.

Although the significant results presented for conversion analysis were limited by sample sizes, it is likely that larger sample sizes would reinforce the lower conversion rates as the expressed interest level drops. The effects by gender and race are not so apparent and further analysis is recommended.

C. RECOMMENDATIONS

This study has shown that interest in joining the military can be predicted using regionally available independent variables; unemployment, parents' education and intentions of going to college. Furthermore, it is apparent that conversion of that interest to military enlistment actions varies by interest level. It is recommended that the recruiting commands incorporate this information into their

recruiting goal models so that the allocation of required new recruits more accurately reflects regional realities.

A number of future projects continuing this line of work may prove fruitful. First, the propensity estimating equations should be aggregated to the regional level. Now that the predictor variables, their regional proxies and the proper coefficients for interest level estimation have been identified, it is a fairly straightforward task to obtain the necessary regional information and compute expected levels of interest for the specific regions.

Second, one could match the YATS and DMDC data for additional years. With this additional data, a more precise estimation of coefficient values will be possible. This may point out differing effects of the predictor variables by race and gender. Because of the improved accuracy of the coefficients, it may also be possible to determine if separate estimating equations are necessary for differing races and genders. The additional sample size would also allow for a more reliable estimation of conversion rates, especially by race and gender as well as by interest level. Finally, additional years of data would provide an opportunity to verify the predictive ability of these propensity estimating equations with data not used in their estimation.

Finally, one could consider other possible predictors of interest that may not be directly available on a regional level, but that could be obtained. There are some

possibilities for predictor variables that, although a direct regional proxy is not available, highly correlated regional data may be available. Examples could be college education costs, the number of military relatives or acquaintances, or the quality of education.

The other area for future research touched on by this analysis is a continued look at predicting mental aptitude. The equation developed in this research predicted actual AFQT percentile. It was then used to eliminate those not expected to score above the fiftieth percentile. With some minor modification, the selection procedure could be used to analyze propensity for differing levels of mental aptitude. A suggestion would be to segregate the sample by mental group category (I-V) and then estimate military interest by mental group.

This research may prove beneficial to the military recruiting commands. Combined with the recommended follow on work, it may be directly implementable for assigning recruiting goals to the services' recruiting areas. Although interest appears to be predictable, it must be combined with conversion rate information to provide an accurate measure of the recruit market in an area. Military recruiting is a form of warfare; the goal, however is to 'capture' the opposition rather than destroy it. This research has provided a start for developing a needed sensor to find the potential recruits.

APPENDIX A

TABLE 1

Cross-Classification of MEPS Codes and Service Specific Recruiting Area Codes

Name	MEPS Code	Army Code	Navy Code	Marine Corps Code	Air Force Code
Portland, ME	01	1D	102	1971	19
Manchester, NH	02	1A	102	1971	19
Boston, MA	03	1C	102	1930	12
Springfield, MA	04	1F	101	1950	12
New Haven, CT	05	1H	101	1950	12
Albany, NY	06	1A	101	1922	16
Fort Hamilton, NY	07	1G	104	1980	14
Newark, NJ	08	1I	161	1979	16
Philadelphia, PA	09	1K	119	4986	15
Syracuse, NY	10	1N	103	1922	13
Buffalo, NY	11	1N	103	1932	13
Wilkes-Barre, PA	12	1E	105	4987	18
Harrisburg, PA	13	1E	105	4987	18
Pittsburg, PA	14	1L	420	4988	11
Baltimore, MD	15	1B	409	4926	35
Richmond, VA	16	3K	408	4994	34
Beckley, WV	17	3B	407	4934	34
Knoxville, TN	18	3I	314	6976	32
Nashville, TN	19	3I	314	6976	32
Louisville, KY	20	3F	407	4968	32
Cincinnati, OH	21	5B	418	4938	52
Columbus, OH	22	5D	418	4938	52
Cleveland, OH	23	5C	417	4940	53
Detroit, MI	24	5F	422	9963	54
Milwaukee, WI	25	5J	559	9974	55

Cross-Classification of MEPS Codes and Service Specific Recruiting Area Codes					
Name	MEPS Code	Army Code	Navy Code	Marine Corps Code	Air Force Code
Chicago, IL	26	5A	521	9936	51
Indianapolis, IN	27	5H	423	9956	50
St. Louis, MO	28	5N	524	9804	45
Memphis, TN	29	4F	347	6976	48
Jackson, MS	30	4F	347	6928	48
New Orleans, LA	31	4I	734	8978	46
Montgomery, AL	32	3H	310	6928	31
Atlanta, GA	33	3A	313	6970	31
Fort Jackson, SC	34	3D	311	6970	37
Jacksonville, FL	35	3E	312	6960	33
Miami, FL	36	3G	348	6960	33
Charlotte, NC	37	3C	315	6992	37
Raleigh, NC	38	3J	315	6992	37
Shreveport, LA	39	4H	734	8978	44
Dallas, TX	40	4C	731	8942	44
Houston, TX	41	4E	732	8952	46
San Antonio, TX	42	4K	746	8998	41
Oklahoma City, OK	43	4J	733	8982	49
Amarillo, TX	44	4J	730	8924	67
Little Rock, AR	45	4H	733	8964	48
Kansas City, MO	46	4G	527	8962	49
Des Moines, IA	47	5E	529	9946	43
Minneapolis, MN	48	5K	528	9972	56
Fargo, ND	49	5L	528	9972	56
Sioux Falls, SD	50	5L	529	8984	43
Omaha, NE	51	5L	529	8984	43

Cross-Classification of MEPS Codes and Service Specific Recruiting Area Codes					
Name	MEPS Code	Army Code	Navy Code	Marine Corps Code	Air Force Code
Denver, CO	52	4D	725	8944	67
Albuquerque, NM	53	4A	730	8924	67
El Paso, TX	54	4A	730	8924	41
Phoenix, AZ	55	6G	840	12989	62
Salt Lake City, UT	56	6J	837	12989	68
Butte, MO	57	6J	839	12802	68
Spokane, WA	58	6L	839	12802	68
Boise, ID	59	6J	837	12990	68
Seattle, WA	60	6L	839	12802	61
Portland, OR	61	6H	837	12990	61
Oakland, CA	62	6I	838	12995	66
Fresno, CA	63	6A	838	12995	63
Los Angeles, CA	64	6F	836	12966	69
San Diego, CA	68	6K	840	12999	62
Tampa, FL	69	3E	312	6960	33
Total	66	52	41	44	35

SOURCE: Extracted from Sampling Design and Sample Selection Procedures, Youth Attitude Tracking Study, 1988 by Immerman et al.

TABLE 2

Sample Size (Interviews) by MEPS based on 1988 Allocation[1]					
Name	MEPS Code	Young Males	Older Males	Young Females	Older Females
Portland, ME	01	73.6	11.7	20.5	9.2
Manchester, NH	02	28.8	5.5	17.6	4.4
Boston, MA	03	81.6	21.8	66.5	24.9
Springfield, MA	04	80.9	11.9	26.1	7.2
New Haven, CT	05	90.8	16.8	26.3	6.5
Albany, NY	06	54.8	9.6	20.8	10.0
Fort Hamilton, NY	07	172.6	41.7	125.7	39.3
Newark, NJ	08	95.8	18.8	92.7	28.4
Philadelphia, PA	09	78.0	15.1	68.7	22.4
Syracuse, NY	10	29.8	7.0	22.4	6.8
Buffalo, NY	11	85.1	13.5	35.4	6.9
Wilkes-Barre, PA	12	35.5	8.7	22.6	5.5
Harrisburg, PA	13	44.3	10.6	33.7	12.2
Pittsburg, PA	14	83.6	19.2	54.6	24.8
Baltimore, MD	15	76.1	17.6	75.5	18.6
Richmond, VA	16	84.2	19.7	40.6	15.4
Beckley, WV	17	99.0	13.1	20.4	7.2
Knowville, TN	18	43.8	11.0	34.2	11.9
Nashville, TN	19	47.7	11.3	37.5	13.2
Louisville, KY	20	94.1	13.2	45.9	11.2
Cincinnati, OH	21	91.5	18.6	43.2	15.9
Columbus, OH	22	91.6	19.4	47.5	15.2
Cleveland, OH	23	78.0	19.3	74.7	19.2
Detroit, MI	24	112.0	21.5	112.5	33.7
Milwaukee, WI	25	85.9	21.1	61.4	22.3
Chicago, IL	26	152.6	33.9	139.8	37.9

Sample Size (Interviews) by MEPS based on 1988 Allocation[1]					
Name	MEPS Code	Young Males	Older Males	Young Females	Older Females
Indianapolis, IN	27	78.5	13.5	52.5	18.3
St. Louis, MO	28	88.1	15.4	55.0	23.5
Memphis, TN	29	52.2	12.5	34.8	8.6
Jackson, MS	30	36.9	8.5	24.0	8.0
New Orleans, LA	31	98.9	19.6	47.7	14.7
Montgomery, AL	32	97.9	22.3	60.6	20.4
Atlanta, GA	33	92.7	17.8	64.6	25.3
Fort Jackson, SC	34	98.1	19.1	53.9	13.0
Jacksonville, FL	35	48.6	10.6	33.4	11.7
Miami, FL	36	102.9	28.0	68.6	27.1
Charlotte, NC	37	93.7	17.2	51.8	14.2
Raleigh, NC	38	95.3	16.4	38.4	16.6
Shreveport, LA	39	26.7	5.1	18.9	7.9
Dallas, TX	40	81.9	19.5	69.1	23.6
Houston, TX	41	90.1	21.0	62.0	19.1
San Antonio, TX	42	93.5	15.4	57.6	19.5
Oklahoma City, OK	43	89.9	15.8	44.4	17.9
Amarillo, TX	44	16.2	3.1	11.8	4.9
Little Rock, AR	45	88.9	13.1	29.0	8.5
Kansas City, MO	46	73.6	14.6	49.9	16.6
Des Moines, IA	47	74.2	15.4	40.4	13.1
Minneapolis, MN	48	78.9	19.4	47.0	23.0
Fargo, ND	49	12.8	2.7	11.4	3.5
Sioux Falls, SD	50	31.3	6.1	11.8	6.0
Omaha, NE	51	55.8	9.2	20.2	6.5
Denver, CO	52	85.4	15.6	50.3	19.1
Albuquerque, NM	53	40.3	7.6	11.4	5.3
El Paso, TX	54	47.8	9.5	15.0	4.8

Sample Size (Interviews) by MEPS based on 1988 Allocation[1]					
Name	MEPS Code	Young Males	Older Males	Young Females	Older Females
Phoenix, AZ	55	94.2	19.0	45.3	13.7
Salt Lake City, UT	56	46.9	9.6	23.8	6.6
Butte, MO	57	19.8	4.6	10.2	3.1
Spokane, WA	58	19.2	4.2	16.9	6.1
Boise, ID	59	21.6	4.7	10.4	4.4
Seattle, WA	60	58.6	10.5	34.8	17.6
Portland, OR	61	96.1	18.1	33.6	12.7
Oakland, CA	62	107.6	25.1	99.6	31.0
Fresno, CA	63	99.3	15.3	30.1	10.8
Los Angeles, CA	64	220.4	52.1	146.4	49.1
San Diego, CA	68	91.4	13.7	27.7	13.5
Tampa, FL	69	62.6	15.3	48.9	13.7
Total		5000.5	1017.7	3030.2	1013.1

[1]These values are fractional because they are the expected values as determined by historic interview completion rates and the necessary MEPS and market segment distribution requirements.

SOURCE: Extracted from Sampling Design and Sample Selection Procedures, Youth Attitude Tracking Study, 1988 by Immerman et al.

TABLE 3

Precision Requirements for the 1988 Sample	
Population	Required Precision[1]
1. Younger Male	
National Level Estimate	0.0100
Recruiting Area Estimate Based on Total Area Population	
less then 100,000	0.0750
100,000-149,999	0.0750
150,000-199,999	0.0750
200,000-249,999	0.0650
250,000-299,999	0.0550
300,000-349,999	0.0500
greater then 349,999	0.0287
2. Older Male	
National Level Estimate	0.0175
3. Younger Female	
National Level Estimate	0.0102
4. Older Female	
National Level Estimate	0.0175

[1]The precision is in terms of the maximum value of the standard error associated with the estimation of the proportion of people with a propensity for enlistment.

SOURCE: Extracted from Sampling Design and Sample Selection Procedures, Youth Attitude Tracking Study, 1988 by Immerman et al.

APPENDIX B

These are the questions asked in the YATS survey that are referenced in this research.

Q406- Do you have a regular high school diploma, a GED, an ABE, or some other kind of certificate (of high school completion)?

01 = REGULAR HIGH DIPLOMA
02 = ABE
03 = GED
04 = SOME OTHER KIND OF CERTIFICATE OF HIGH SCHOOL EQUIVALENCY
05 = NONE OF THE ABOVE

Q407- (In October, will you be/Are you) enrolled in any school, college, vocational or technical program, apprenticeship, or job training course?

01 = YES
02 = NO

Q408- What kind of school or training program (will you be/are you) enrolled in?

01 = NO SCHOOLS OR TRAINING PROGRAMS
02 = ADULT BASIC EDUCATION (ABE)
03 = TAKING HIGH SCHOOL CLASSES IN A REGULAR, DAY HIGH SCHOOL
04 = GED OR H.S. EQUIVALENCY PROGRAM
05 = SKILL DEVELOPMENT PROGRAM
06 = ON-THE-JOB TRAINING PROGRAM
07 = APPRENTICESHIP PROGRAM
08 = VOCATIONAL, BUSINESS, OR TRADE SCHOOL
09 = 2-YEAR JUNIOR OR COMMUNITY COLLEGE
10 = 4-YEAR COLLEGE OR UNIVERSITY

Q408A- This is the same as Q408.

Q436- How easy or difficult is it for someone of your age to get a full time job in your community? Is it...

01 = ALMOST IMPOSSIBLE
02 = VERY DIFFICULT
03 = SOMEWHAT DIFFICULT, OR
04 = NOT DIFFICULT AT ALL?

Q437- And how easy or difficult is it for someone your age to get a part-time job in your community? Is it...
01 = ALMOST IMPOSSIBLE
02 = VERY DIFFICULT
03 = SOMEWHAT DIFFICULT, OR
04 = NOT DIFFICULT AT ALL?

Q503- How likely is it that you will be serving in the military? Would you say...
01 = DEFINITELY
02 = PROBABLY
03 = PROBABLY NOT, OR
04 = DEFINITELY NOT?

Q517- We've talked about several things you might be doing in the next few years. Taking everything into consideration, what are you <u>most likely</u> to be doing (in October 198X--that is, a year from this fall/after you finish high school)?
01 = GOING TO SCHOOL FULL TIME
02 = GOING TO SCHOOL PART TIME
03 = WORKING FULL-TIME
04 = WORKING PART-TIME
05 = SERVING IN THE MILITARY
06 = BEING A FULL-TIME HOMEMAKER
07 = OTHER

Q522- Now, I'd like to ask you in another way about the likelihood of your serving in the military. Think of a scale from zero to ten, with ten standing for the *very highest* likelihood of serving and zero standing for the very lowest likelihood of serving. How likely is it that you will be serving in the military in the next few years?
01 = 01 etc.

Q622- Within the last 12 months, have you made a toll-free call for information about the military?
01 = YES
02 = NO

Q625- Within the last 12 months have you sent a postcard or coupon for information about the military?
01 = YES
02 = NO

Q628- Have you <u>ever</u> talked with any military recruiter to get information about the military?
01 = YES
02 = NO

Q645- Have you ever taken the three hour written test called the ASVAB that is required to enter the military?

01 = YES
02 = NO

Q683- Within the last year or so, have you discussed with anyone the possibility of your serving in the military?

01 = YES
02 = NO

Q692- How do <u>you</u> feel about serving in the active military yourself? Are you...

01 = VERY FAVORABLE
02 = SOMEWHAT FAVORABLE
03 = NEITHER FAVORABLE NOR UNFAVORABLE
04 = SOMEWHAT UNFAVORABLE, OR
05 = VERY UNFAVORABLE?

CPYATS82- COMPOSITE ACTIVE PROPENSITY [Most positive response to the four Service-specific propensity (for active duty) questions]

01 = DEFINITELY
02 = PROBABLY
03 = PROBABLY NOT
04 = DEFINITELY NOT

V438JOIN- Joining the (military/service).

01 = MENTIONED (as a likely pursuit in the next few years)
02 = NOT MENTIONED

SOURCE: Extracted from <u>Youth Attitude Tracking Study II Wave 17--Fall 1986 Codebook</u> by the Research Triangle Institute.

APPENDIX C

This if/then coding was used within SAS to classify the respondents' interest level in joining the military.

```
IF V438JOIN=01 OR Q503=01 OR CPYATS82=01 OR Q522=08 OR
Q522=09 OR Q522=10 THEN INTEREST=1;

ELSE IF Q503=04 AND CPYATS82=04 AND (Q522=00 OR Q522=01 OR
Q522=02 OR Q522=03) THEN INTEREST=2;

ELSE INTEREST=3;
```

LIST OF REFERENCES

Bain, L.J. and Engelhardt, M., Introduction to Probability and Mathematical Statistics, Duxbury Press, 1987.

Borack, J.I., A Framework for Integrating Alternative Military Manpower Supply Methods, Navy Personnel Research and Development Center, San Diego, California, 1984.

Goldfeld, S.M. and Quandt, R.E., Nonlinear Methods in Econometrics, North-Holland Publishing Company, 1972.

Gorman, L. and Mehay, S.L., Estimating Local Area Propensity for the USAR: A Feasibility Analysis, Naval Postgraduate School, Monterey, California, 1989.

Telephone conversation between Major John Hershberger, United States Army Recruiting Command and the author on 31 August 1990.

Huzar, C.E., Profiling Market Potential for Navy Recruiting at the Local Geographic Level, Master's Thesis, Naval Postgraduate School, Monterey, California, 1988.

Immerman, F.W., York, B.J., and Mason, R.E., Sampling Design and Sample Selection Procedures, Youth Attitude Tracking Sturdy, 1988, Research Triangle Institute, 1987.

Judge, G.G., and others, The Theory and Practice of Econometrics, John Wiley and Sons, 1985.

Maddala, G.S., Limited-Dependent and Qualitative Variables in Econometrics, Cambridge University Press, 1983.

Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics), America's Volunteers: A Report on the All-Volunteer Armed Forces, Government Printing Office, Washington D.C., 1978.

Ogle, L.T., and Asalam, N., The Condition of Education Volume 1 Elementary and Secondary Education, U.S. Department of Education Office of Educational Research and Improvement, 1990.

Orvis, B.R., Forecasting Enlistment Actions from Intention Information: Validity and Improvement N-1954-MRAL, The RAND Corporation, 1982.

Orvis, B.R. and Gahart, M.T., Relationship of Enlistment Intention and Market Survey Information to Enlistment in Active Duty Military Service N-2292-MIL, The RAND Corporatio.ı, 1985.

Orvis, B.R. and Gahart, M.T., Quality Based Analysis Capability for National Youth Surveys; Development, Application, and Implications for Policy R-3675-FMP, The RAND Corporation, 1989.

Peterson, J., AFQT Score Forecasting Models for Regional Estimation of Qualified Military Available, Master's Thesis, Naval Postgraduate School, Monterey, California, June 1990.

Waksberg, J., "Sampling Methods for Random Digit Dialing", Journal of the American Statistical Association, v.73, 1978.

INITIAL DISTRIBUTION LIST

1.	Defense Technicl Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 52 Naval Postgraduate School Monterey, California 93943-5002	2
3.	Commander, Code 22 Navy Recruiting Command 4015 Wilson Blvd Arlington, Virginia 22203	2
4.	Professor George Thomas, Code AS/Te Department of Administrative Sciences Naval Postgraduate School Monterey, California 93943	1
5.	Professor Linda Gorman, Code AS/Gr Department of Administrative Sciences Naval Postgraduate School Monterey, California 93943	1
6.	Professor Dan C. Boger, Code OR/Bo Department of Operations Research Naval Postgraduate School Monterey, California 93943	1
7.	Lieutenant Richard Snyder 544 Marshview Dr. Jacksonville Beach, Florida 32250	2
8.	Operations Analysis Programs Code 30 Naval Postgraduate School Monterey, CA 93943	1